Analysis of measured and simulated supraglottal acoustic waves

Rubén Fraile , Vera V. Evdokimova , Karina V. Evgrafova , Juan I. Godino-Llorente , Pavel A. Skrelin

13 **Abstract**

To date, while much attention has been paid to the estimation and modelling of the voice source (i.e. the glottal airflow volume velocity), the measurement and characterisation of the supraglottal pressure wave have been much less studied. Some previous results have unveiled that the supraglottal pressure wave has some spectral resonances similar to those of the voice pressure wave. This makes the supraglottal wave partially intelligible. While the explanation for such effect seems to be clearly related to the reflected pressure wave travelling upstream along the vocal tract, the influence that non-linear source-filter interaction has on it is not as clear. This paper provides an insight into this issue by comparing the acoustic analyses of measured and simulated supraglottal and voice waves. Simulations have been performed using a high-dimensional discrete vocal-fold model. Results of such comparative analysis indicate that spectral resonances in the supraglottal wave are mainly caused by the regressive pressure wave that travels upstream along the vocal tract and not by source-tract interaction. On the contrary and according to simulation results, source-tract interaction has a role in the loss of intelligibility that happens in the supraglottal wave with respect to the voice wave. This loss of intelligibility mainly corresponds to spectral differences around the frequency of the second formant.

Keywords: Speech analysis; Biomedical acoustics; Biomechanical modelling; Voice production modelling & simulation.

1. Introduction

Voice production is frequently modelled as a source-system model [1] in which the larynx acts as a source producing a glottal airflow which is subsequently processed by the vocal and nasal tracts (i.e. the system or filter) to produce a radiated pressure wave at the lips. The glottal airflow, that is, the volume of air that goes through the glottis per time unit (glottal airflow volume velocity) is also named as the voice source, and it is closely related to vocal fold vibration [2]. For this reason, knowing it becomes interesting for many applications of voice analysis. However, since the larynx cannot be separated from the vocal tract, the actual voice source cannot be measured; instead, it has to be estimated. Such estimation can be based on the voice pressure wave measured at the output of the lips (voice wave) [2], on measurements of mouth airflow volume velocity [3] or on combining information obtained from the voice wave and from other signals, such as the electroglottogram [4].

Furthermore, even in the event that the larynx could be somehow separated from the vocal tract, the voice source could not be measured either. This is due to the non-linear interaction that happens between the larynx and the vocal tract [5]. Physically, such interaction is related to the fact that the air pressure variations immediately above the larynx (i.e. the supraglottal pressure wave, or supraglottal wave for short) pose a contour condition to the flow of air across the glottis. Mathematically, this contour condition can be taken into account by considering the value of the regressive pressure in the calculation of the glottal airflow volume velocity (e.g. equation 33 in [6]).

In spite of voice source and supraglottal wave being closely related, while much attention has been paid to the estimation and modelling of the voice source [2], the measurement and characterisation of the supraglottal wave have been much less studied. Cranen and Boves proposed a method for *in vivo* measuring of pressure at different levels along the voice production system [7]. They devised a procedure for the calibration of pressure transducers and, based on their measurements, they concluded that the load impedances posed by vocal tract and trachea on the glottis had an effect on the skewing of the voice source. A simpler but similar procedure based on using nasal catheters for introducing the transducers in the pharynx had been previously used by Lisker to characterise the relation between supraglottal and intraoral pressures in the production of English stops [8]. He concluded that articulation, particularly the production of stops in /p,t,k/ and /b,d,g/, was not independent from supraglottal pressure.

More recently, Evdokimova *et al* used a microphone, also mounted on a nasal catheter, to record the supraglottal wave synchronously to an outer microphone measuring the voice wave [9]. Their results indicated that the spectral resonances present on the supraglottal wave were affected by articulation in the case of Russian vowels. In [10], Evgrafova *et al* reported on a perceptual experiment carried out with the same supraglottal and voice wave recordings as in [9]. The analysis of those results confirmed that the supraglottal wave had some spectral resonances similar to those of the voice wave, and that these had the effect of making the supraglottal wave partially intelligible, at least in what refers to the identification of vowels. According to the conclusions of the perceptual experiments reported by Dubno and Dorman [11], such a limited loss of intelligibility in vowels should be related to the first formant remaining unaltered to a great extent.

While *a priori* the explanation for the presence of spectral resonances on the supraglottal wave can be attributed to the effect of the reflected pressure wave travelling upstream along the vocal tract, it presently remains unknown whether non-linear source-filter interaction plays a relevant role in this event. In this paper, an insight into this issue is provided by comparing the acoustic analyses of measured supraglottal and voice waves with results obtained from simulating these signals using the discrete multiple-mass simulator described in [12].

2. Materials

2.1. Acoustic measurements

Voice recordings from a male and a female, both native Russian speakers, are available. Recordings from the female speaker correspond to four utterances of each one of the 6 Russian vowels: /a,e,i,ɨ,o,u/. For the male speaker, only two utterances per vowel are available. For each utterance, the voice wave was measured at the output of the lips using a head-mounted microphone (AKG HSC20). Simultaneously, the supraglottal wave was measured by inserting a miniature waterproof microphone (QueAudio, 2.3 mm diameter) through the nasal cavity. All recordings were sampled at 44,100 Hz and had a precision of 16 bits [9, 10].

Figure 1 shows a pair of synchronously measured supraglottal and voice waves corresponding to the female speaker. Two significant features can be noticed in the supraglottal signal. Firstly, it has a low frequency component, somewhat of a baseline drift, probably a motion artefact due to the miniature microphone not being at a fixed position relative to the centre of the larynx. Secondly, the signal waveform has a sort of saturation which implies clipping of its most prominent peaks, be them either positive or negative. This is due to the fact that the maximum sound pressure level (SPL) that can be measured by the microphone is 130 dB (63 Pa), while for typical phonatory set-ups the supraglottal pressure reaches peak pressure values in the vicinity of 6 cmH_2O [7, 8], i.e. approximately 590 Pa. The simulated supraglottal wave shown in Figure 2 has lower peak values (182 Pa), but its root mean square value is in the range of 100 Pa, which corresponds to SPL $\approx$ 134 dB, still beyond the dynamic range of the microphone. Figure 1 shows that for moderate amplitudes of the voice wave only the greatest peaks of the supraglottal wave are affected by this kind of distortion, while for higher amplitudes less significant peaks become affected too. Such peak-clipping has little impact on speech intelligibility and its main spectral effect is a flattening of the signal's spectrum. A deeper insight into this issue is provided in the Appendix.

2.2. Perceptual tests

The recordings corresponding to the supraglottal waves were played for 25 listeners, among which 5 were experts in phonetics and 20 were naive listeners [10]. The expert listeners correctly recognised almost all the vowels, except for some confusions between /i/ and /ɨ/. The naive listeners reported acoustic

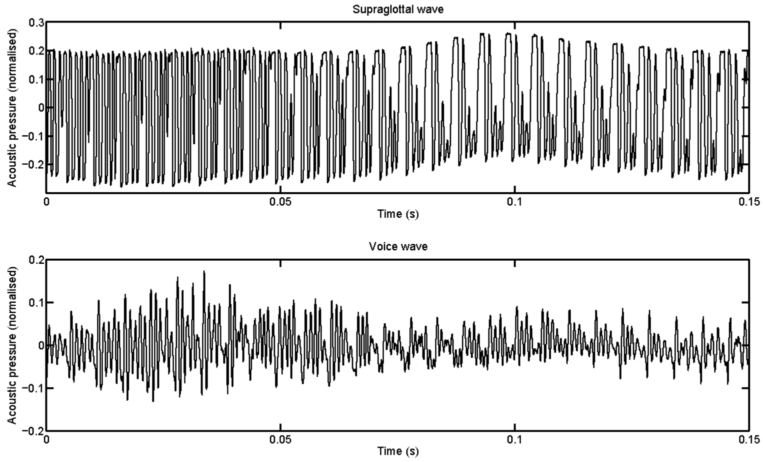

Figure 1: Acoustic wave recordings corresponding to one phonation of vowel /a/ performed by the female speaker: supraglottal (top) and voice (bottom) waves.

differences in the sounds corresponding to different vowels and they were able to correctly identify vowels /a/, /e/ and /i/ in the majority of cases. However, there were frequent confusions between /i/, /ɨ/ and /u/, and also between /e/ and /o/.

3. Methods

3.1. Simulation

The voice production simulation model described in [12] was used to simulate the production of one open vowel (/ɑ/), one closed front vowel (/i/), and one closed back vowel (/u/). The vocal tract area functions used for simulation were the ones corresponding to individual TB in [13]. For each vowel, two simulation modes were implemented: coupled and uncoupled. In the coupled mode, the regressive pressure wave p_{01}^{-}, coming upwards from the vocal tract, affects the voice source. This is modelled by making the calculation of the glottal airflow volume velocity u_{g} dependant on the magnitude of such regressive wave. This dependency has the form indicated by equation (23) in [12]:

$$u_{\mathrm{g}} = c\frac{a_{\mathrm{g}}}{k_{\mathrm{t}}} \cdot \left(-\frac{a_{\mathrm{g}}}{a} \pm \sqrt{\left(\frac{a_{\mathrm{g}}}{a}\right)^2 + \frac{4k_{\mathrm{t}}\left|p_{\mathrm{sub}}^{+} - p_{01}^{-}\right|}{c^2\rho_0}} \right) \tag{1}$$

In the uncoupled mode the glottal flow is made independent of the regressive pressure waves coming upstream from the vocal tract. This is achieved by making the variable p_{01}^{-} equal to 0. Conceptually, this means that the simulated larynx behaves as if its outgoing airflow was transmitted to an infinitely wide

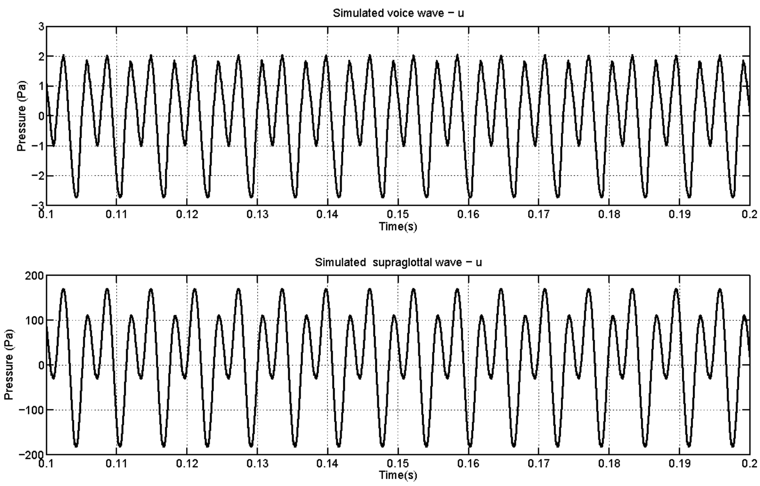

Figure 2: Voice wave(top) and supraglottal wave (bottom) obtained from simulation of vowel /u/ in coupled mode.

open space. In this set-up, the vocal tract does not interact with the larynx and the voice source only depends on the subglottal pressure p^{+}_{sub} and some geometrical parameters describing the configuration of the vocal folds. Equation (23) in [12] is simplified to the following form:

$$u_{\text{g}} = c\frac{a_{\text{g}}}{k_{\text{t}}} \cdot \left(-\frac{a_{\text{g}}}{a} \pm \sqrt{\left(\frac{a_{\text{g}}}{a}\right)^2 + \frac{4k_{\text{t}}\left|p^{+}_{\text{sub}}\right|}{c^2\rho_0}} \right) \tag{2}$$

Independence between glottal flow and vocal tract is a key assumption on which the source-filter model of voice production is grounded [1].

Figure 2 shows one fragment of a voice wave simulated in the coupled mode with its corresponding supraglottal wave.

3.2. Acoustic signal processing

Each pair of recordings corresponding to the same utterance has been time-aligned to compensate for the delay in the propagation of the acoustic wave from the larynx to the lips and to the outer microphone. Alignment was performed by looking for the maximum in the cross-correlation between both signals. After that, a frame having a duration of 0.5 s was extracted from each pair of recordings, corresponding to the central part of the utterance.

For the sake of comparison of signals in spectral domain, the spectral envelope of each signal was estimated using the non-parametric Blackman-Tukey method [14, p.879] with a frequency resolution of 75 Hz. The Blackman-Tukey method is based on calculating the Fourier transform of the estimated autocorrelation of the signal. Setting a resolution in spectral domain to 75 Hz implies that the autocorrelation has to be estimated only for time lags shorter than $\frac{1}{75} \approx 13.3$ ms. For a signal duration equal to 0.5 s, this means keeping only the autocorrelation estimates with the lowest variances, thus discarding estimates with the highest variances that correspond to the longest time lags [15, p.569].

Formant positions were estimated for voice and supraglottal waves using the algorithm described in [16]. Formants not correctly estimated were discarded after visual inspection.

4. Results

4.1. Analysis of simulated glottal flow

Figure 3 depicts the glottal flows resulting from the simulation of the production of vowels /ɑ/ and /i/, in both the coupled and the uncoupled modes. In addition to the time-domain waveforms, their spectrum estimates have been included in the figure. The two dominant effects of the coupling between the larynx and the vocal tract are the increased skewness on the glottal flow waveform and a reduction in the fundamental frequency. The increased skewness is consistent with the findings of Childers and Wong [17] after a joint analysis of the voice source, estimated by inverse filtering, and the electroglottographic signal. The reduced fundamental frequency is in agreement with the results of the theoretical analysis and simulations reported by Titze [5]. The fact that the effects of source-tract interaction were more relevant in the case of closed vowels and narrower epilarynges [5] is consistent with the higher difference in fundamental frequency observed for /i/ (right) with respect to /ɑ/ (left).

As for the spectrum estimates, no relevant differences can be noticed between the coupled and uncoupled modes. As a consequence, the presence of resonances in the supraglottal acoustic wave does not seem to be attributable to the effect of source-tract interaction on the voice source. In other words, the glottal flow does not exhibit spectral resonances, even when simulations are run in coupled mode. This implies that resonances present in the supraglottal wave are not caused by corresponding resonances in the voice source.

4.2. Analysis of the simulated acoustic signals

Figure 4 shows the spectrum estimates of the acoustic signals, both the supraglottal and the voice waves, corresponding to the simulated vowels /ɑ/, /i/, and /u/. In general terms, the spectral tilt is very similar for the supraglottal and the voice waves, both in the coupled and the uncoupled modes of simulation. This is coherent with the fact that spectral tilt is mainly affected by the voice source (Figure 3) and not by the propagation of the pressure wave along the vocal tract. Another common feature of all results, except for /ɑ/ in the uncoupled mode, is that the spectrum on the supraglottal wave becomes flatter than that of the voice wave at frequencies above 2,000 Hz. Moreover, except for /u/, the supraglottal wave spectrum is flatter for the coupled mode than for the uncoupled mode.

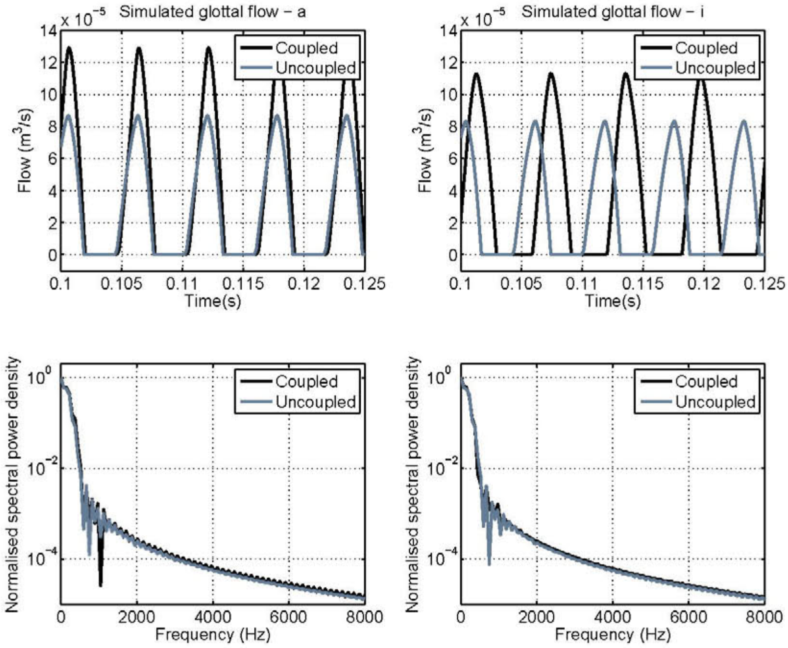

Figure 3: Glottal flow corresponding to the simulation of vowels /ɑ/ (left) and /i/ (right), both in the coupled and uncoupled modes: waveforms (top) and their corresponding estimated spectra (bottom) calculated using the Blackman-Tukey method.

From the point of view of the intelligibility of vowels, the positions of the first two formants are critical. The simulation results for /i/ and /u/ show that the spectra of the supraglottal and voice waves is very similar for frequencies below 1,000 Hz. Therefore, the first formant (F1) is similar for both signals. In the case of /ɑ/, the low-frequency part of the spectrum is more different for both signals. However, the supraglottal wave spectrum still has its highest peak near the highest resonance of the voice wave, although the precise frequency at which the peak is placed may be different. As for the second formant (F2), whose frequency typically ranges approximately from 900 Hz to 2,500 Hz, it lies in a frequency interval for which the differences the supraglottal and voice spectra are relevant. As a consequence, F2 may be shifted or even disappeared in the supraglottal spectrum. Resonances above 2,000 Hz are more attenuated in the supraglottal wave for the coupled mode of simulation (left graphs in Figure 4). This is likely to affect F2 in the case of /ɑ/ and /i/, while in the case of /u/ F2 should be less affected, since it happens at lower frequencies.

The presence of peaks or resonances in the spectral envelope of the supraglottal signal, especially for frequencies around F1, is related to the fact that listeners can identify vowels when hearing it. The peaks in the supraglottal spec-

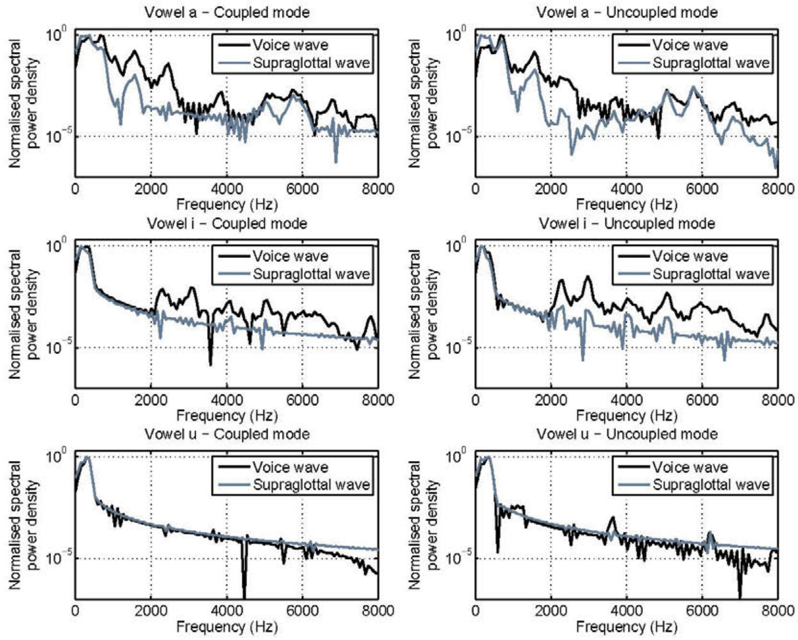

Figure 4: Spectrum estimates of the simulated voice and supraglottal waves corresponding to vowels /ɑ/ (top), /i/ (middle), and /u/ (bottom), both in the coupled (left) and uncoupled (right) modes.

trum simulated using the uncoupled mode indicate that these resonances are not due to the interaction between glottal source and vocal tract, but to the regressive wave coming upstream the vocal tract which sums to the progressive wave coming from the larynx to result in the supraglottal pressure wave. However, the simulation results also indicate that the position of F2 in the supraglottal wave can be more different from that of the voice wave for the coupled mode of simulation. This implies that the non-linear source-tract interaction negatively affects the intelligibility of the supraglottal wave.

From another point of view, the spectral flattening that occurs at high frequencies in the supraglottal wave with respect to the voice wave may also contribute to the loss of intelligibility. Yet, this effect is of secondary importance when compared to the distortions in F1 and, to a greater extent, in F2.

4.3. Analysis of measurements

Figure 5 depicts the estimated spectra for the recorded supraglottal and voice signals corresponding to vowels /a/, /i/, and /u/ and to the female speaker. The spectral flattening effect observed in Figure 4 can only be noticed here in some cases; therefore it is to be discarded as a plausible cause of the perceptual effects

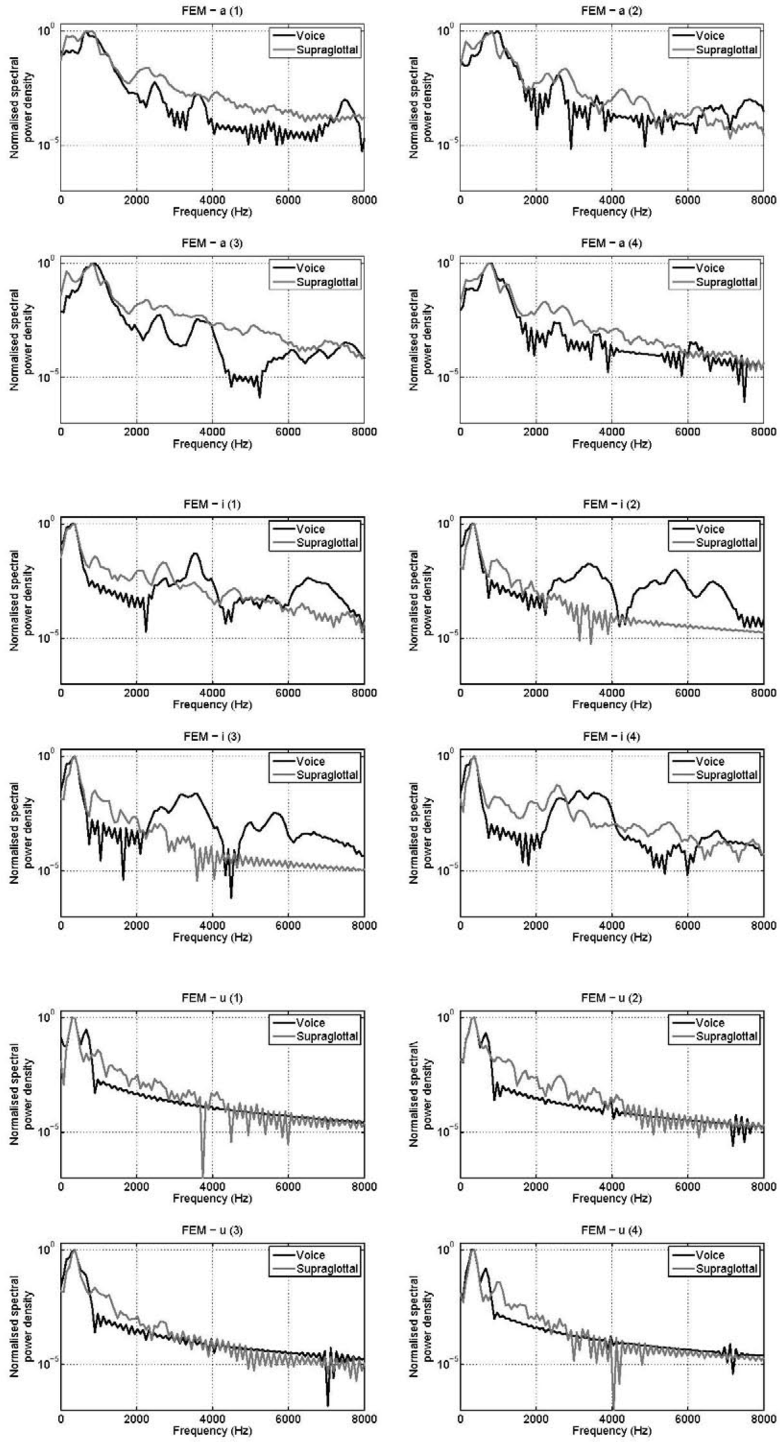

Figure 5: Spectrum estimates of the measured female voice and supraglottal waves corresponding to vowels /a/ (top), /i/ (middle), and /u/ (bottom).

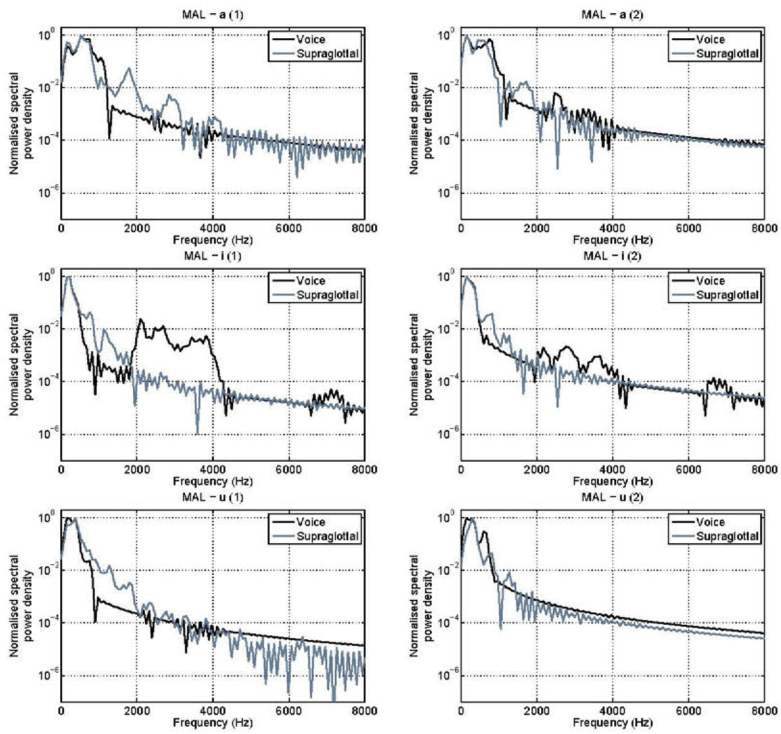

Figure 6: Spectrum estimates of the measured male voice and supraglottal waves corresponding to vowels /a/ (top), /i/ (middle), and /u/ (bottom).

mentioned in section 2.2. Furthermore, this also suggests that the impact of peak clipping in the supraglottal wave is not relevant for this analysis. Regarding the positions of F1 and F2 in the supraglottal and voice signals, the same effects as in Figure 4 can be observed in Figure 5:

- The spectrum around F1 (below 1,000 Hz) is almost coincident for the supraglottal and the voice signals in the case of /i/ and /u/, while the differences are noteworthy in the case of /a/.
- For both vowels, noticeable differences between both spectra happen above 1,000 Hz (2,000 Hz in the case of /a/), hence affecting the position of F2.

A similar behaviour can be observed in the case of the male speaker (Figure 6), although the dissimilarities between supraglottal and voice spectra at low frequencies for vowel /a/ are less remarkable than in the case of the female speaker and the higher-frequency discrepancies happen above 1,000 Hz for all three vowels.

Fig. 7 shows the difference between supraglottal and voice waves of the female speaker in terms of the positions of F1 and F2. The same plot corresponding to the male speaker in in Fig. 8. Overall, it can be appreciated that the position of F2 changes much more than the position of F1, especially in the case of closed vowels (/i/, /ɨ/ and /u/). Furthermore, the positions of these vowels in the F1-F2 plane become very close for the supraglottal wave. This can well explain the difficulties of the listeners to discriminate among these vowels. The same effect can be observed in the case of /e/ and /o/ but their positions stay more separated, which means a higher degree of intelligibility. Last, some recordings corresponding to vowel /a/ experience the greatest changes in F1 from the voice to the supraglottal signals. But the magnitude of such changes is not enough for the values of F1 to become overlapped with those corresponding to /e/.

5. Conclusions

A comparative study of the supraglottal and voice pressure waves has been presented in this paper. Both simulated and measured signals have been analysed with the purpose of finding an explanation for the perceived similarities and differences between both waves.

Listeners find that diverse phonemes can be discriminated when listening to the supraglottal wave, although some mistakes may be made in their identification. The discrimination among vowel phonemes is related to the spectral resonances that are present in the supraglottal wave. However, the spectrum of the simulated voice source does not exhibit such resonances, neither when source-tract interaction is modelled, nor when it is not modelled. These results suggest the conclusion that spectral resonances in the supraglottal wave are mainly caused by the regressive pressure wave that travels upstream along the vocal tract.

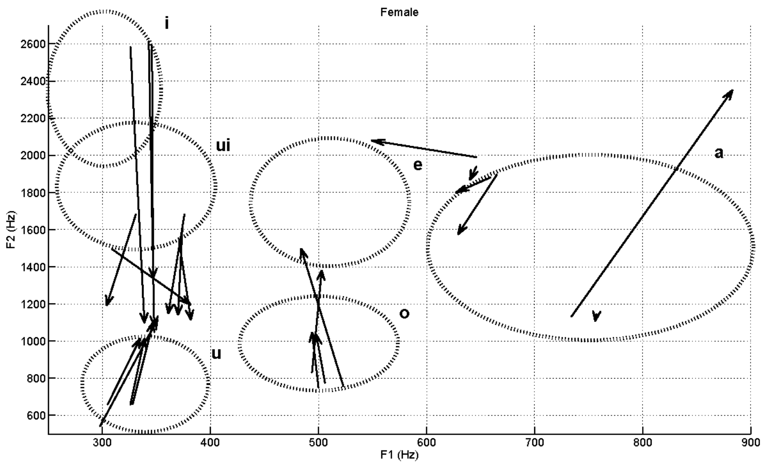

Figure 7: Differences in the positions of F1 and F2 corresponding to the six vowels uttered by the female speaker. Arrows begin at the points corresponding to the voice signal and have their head pointing to the points corresponding to the supraglottal wave. Ellipses indicate the approximate position of Russian vowels in the F1-F2 chart, according to Fig. 2 in [18]. The sign "ui" corresponds to vowel /ɨ/.

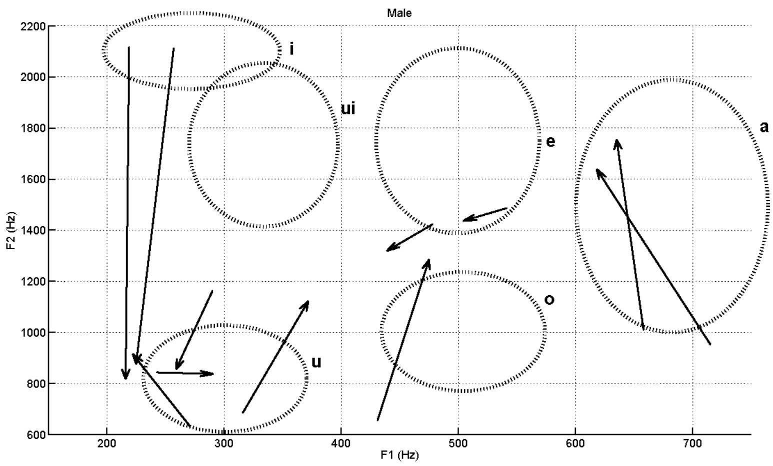

Figure 8: Differences in the positions of F1 and F2 corresponding to the six vowels uttered by the male speaker. Arrows begin at the points corresponding to the voice signal and have their head pointing to the points corresponding to the supraglottal wave. Ellipses indicate the approximate position of Russian vowels in the F1-F2 chart, according to Fig. 2 in [18]. The sign "ui" corresponds to vowel /ɨ/.

Although listeners can discriminate among different phonemes, there is a relevant degree of confusion among /i/, /ɨ/ and /u/ and, to a lesser extent, between /e/ and /o/. This is related to the differences that happen between supraglottal and voice waves in the frequency of F2. Such differences have been consistently identified in measured and simulated voices, so the probability that they have been significantly affected by the measurement process is low. Furthermore, simulation results indicate that the spectral differences in the frequency range corresponding to F2 are enhanced by source-tract interaction.

While some differences in F1 also happen, these are more relevant for open vowels, which are the least confused ones. Therefore, differences in F1 between supraglottal and voice waves are not likely to be the main cause of the reduced intelligibility. Moreover, the fact that listeners can distinguish between front vowels when hearing the supraglottal wave and the findings reported by Dubno and Dorman [11] also suggest that changes in F1 are not relevant enough in this case.

Acknowledgements

This work has been partially financed by the Spanish Government, through project grant number TEC2012-38630-C04-01 ("Evaluación Multimodal de Trastornos Neurológicos mediante la Caracterización de la Voz, Dinámica de los Pliegues Vocales y Secuencias Sacádicas"), and it has also been carried out in the framework of SPbSU project n. 31.37.353.2015 ("The Phonetic Aspects of Speech Signal Synthesis with a High Degree of Naturalness").

Appendix A. Spectral effect of signal clipping due to microphone input saturation

Peak clipping is a non-linear effect that happens when the input signal of a microphone has a dynamic range larger than that of the microphone itself. Due to the non-linear nature of peak clipping, it is not possible to associate it to a given frequency response. On the contrary, its effects on signal's frequency components can only be inferred from either simulation or measurement results. For instance, experiments carried out by Licklider and Pollack [19] indicate that peak clipping affects sound in terms of quality and timber but that intelligibility remains only moderately affected, even in the extreme case in which the signal is transformed into a bivalued (± 1) square signal. Rabiner and Schafer state that peak clipping has the effect of flattening the spectral envelope of the signal [1, p.150]. This effect has been exploited by some researchers for the spectral analysis of voice signals [20].

In order to have an insight into the spectral effect of peak clipping on the specific set of signals used in this paper, a peak-clipping operation has been applied to a measured voice wave $v(t)$. The peak-clipping operation corresponds to the following equation:

$$v_{\text{clip}}(t) = \tanh\left(\frac{v(t)}{\theta}\right) \cdot \theta \tag{A.1}$$

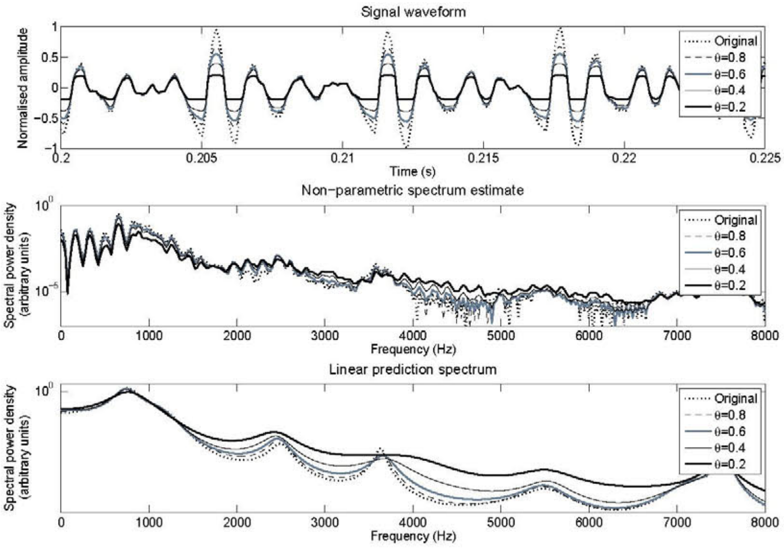

Figure A.9: Effects of peak clipping on signal waveform (top) and on its estimated spectrum via non-parametric (middle) and parametric (bottom) algorithms. The thick line corresponds to the original signal $v(t)$ ($V(f)$ in spectral domain). For the distorted signals, the value of θ in A.1 is specified in the legend. Non-parametric estimation is according to the Blackman-Tukey method [14, p.879] with a frequency resolution of 25 Hz. Parametric estimation corresponds to discrete all-pole modelling [21] with model order 45.

Figure A.9 shows the effects of applying the clipping operation A.1 in spectral domain. As the value of θ departs from 1, the spectral envelope tends to flatten. This effect is more easily noticeable in the parametric estimation of the spectrum (bottom graph), especially for $\theta = 0.2$. However, the resonant frequencies below 3000 Hz remain almost unchanged, which is coherent with the expectancy that intelligibility is not greatly affected by peak-clipping. Therefore, the effect of peak-clipping in this case is as expected: spectrum flattens but resonant frequencies remain stable.

References

[1] L. R. Rabiner, R. W. Schafer, Digital processing of speech signals, Prentice-Hall, 1978.

[2] P. Alku, Glottal inverse filtering analysis of human voice production - A review of estimation and parameterization methods of the glottal excitation and their applications, Sadhana 36 (2011) 623–650.

[3] M. Rothenberg, A new inverse-filtering technique for deriving the glottal air flow waveform during voicing, Journal of the Acoustical Society of America 53 (1973) 1632–1645.

[4] D. E. Veeneman, S. L. BeMent, Automatic glottal inverse filtering from speech and electroglottographic signals, IEEE Transactions on Acoustics, Speech and Signal Processing 33 (1985) 369–377.

[5] I. Titze, Nonlinear source–filter coupling in phonation: Theory, Journal of the Acoustical Society of America 123 (2008) 2733–2749.

[6] I. R. Titze, Parameterization of the glottal area, glottal flow, and vocal fold contact area, Journal of the Acoustical Society of America 75 (1984) 570–580.

[7] B. Cranen, L. Boves, Pressure measurements during speech production using semiconductor miniature pressure transducers: Impact on models for speech production, Journal of the Acoustical Society of America 77 (1985) 1543–1551.

[8] L. Lisker, Supraglottal air pressure in the production of English stops, Language and Speech 13 (1970) 215–230.

[9] V. Evdokimova, K. Evgrafova, P. Skrelin, T. Chukaeva, N. Shvalev, Detection of the frequency characteristics of the articulation system with the use of voice source signal recording method, in: M. Železný, I. Habernal, A. Ronzhin (Eds.), Speech and Computer, volume 8113 of *Lecture Notes in Computer Science*, Springer, 2013, pp. 108–115.

[10] K. Evgrafova, V. Evdokimova, P. Skrelin, T. Chukaeva, N. Shvalev, A new technique to record a voice source signal, in: International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications - MAVEBA 2013, Florence, pp. 181–182.

[11] J. R. Dubno, M. F. Dorman, Effects of spectral flattening on vowel identification, Journal of the Acoustical Society of America 82 (1987) 1503–1511.

[12] R. Fraile, M. Kob, J. Godino-Llorente, N. Sáenz-Lechón, V. Osma-Ruiz, J. Gutiérrez-Arriola, Physical simulation of laryngeal disorders using a multiple-mass vocal fold model, Biomedical Signal Processing and Control 7 (2012) 65–78.

[13] T. Baer, J. C. Gore, L. C. Gracco, P. W. Nye, Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels, Journal of the Acoustical Society of America 90 (1991) 799–828.

[14] J. G. Proakis, D. G. Manolakis, Digital signal processing: principles algorithms and applications, Macmillan Publishing Company, 1988.

[15] K. S. Shanmugan, A. M. Breipohl, Random signals: Detection, estimation, and data analysis, Wiley, 1988.

[16] V. Evdokimova, The use of dynamic vocal tract model for constructing the formant structure of the vowels, in: International Conference on Speech and Computer - SPECOM 2006, St. Petersburg, pp. 210–214.

[17] D. Childers, C. Wong, Measuring and modeling vocal source-tract interaction, IEEE Transactions on Biomedical Engineering 41 (1994) 663–671.

[18] B. M. Lobanov, Classification of Russian vowels spoken by different speakers, Journal of the Acoustical Society of America 49 (1971) 606–608.

[19] J. Licklider, I. Pollack, Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech, Journal of the Acoustical Society of America 20 (1948) 42–51.

[20] J. Schoentgen, Modulation frequency and modulation level owing to vocal microtremor, Journal of the Acoustical Society of America 112 (2002) 690–700.

[21] A. El-Jaroudi, J. Makhoul, Discrete all-pole modeling, IEEE Transactions on Signal Processing 39 (1991) 411–423.